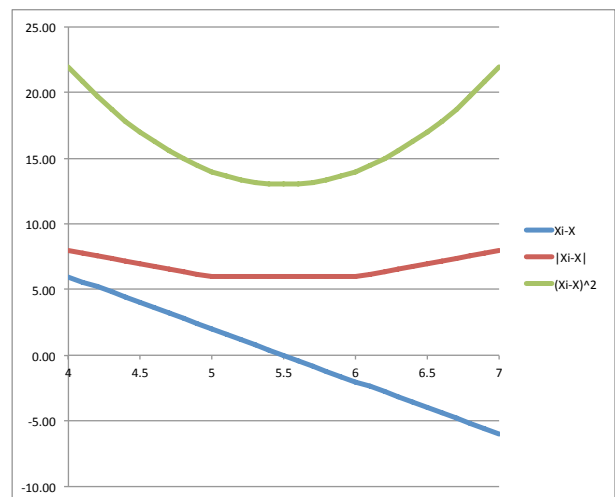
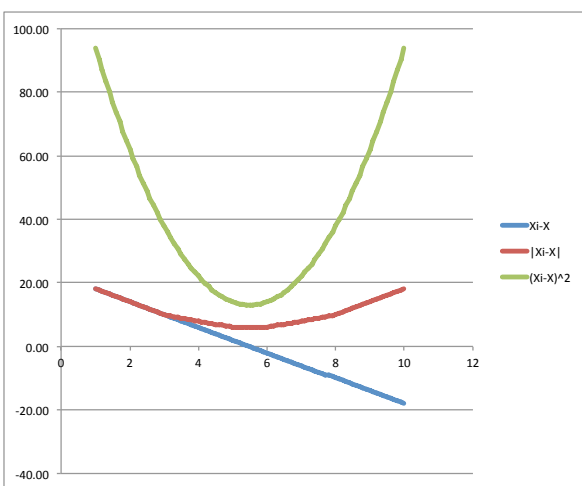


	x_i	$x_i - \bar{x}$	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
x_1	3	-2.5	2.5	6.25
x_2	5	-0.5	0.5	0.25
x_3	6	0.5	0.5	0.25
x_4	8	2.5	2.5	6.25
Σ	22	0	6	13
Mean	5.5	0	1.5	3.25 (/ n) 4.33 (/ n-1)
$\sqrt{\quad}$				1.80 (/ n) 2.08 (/ n-1)

Annahme: Mittelwert = 5

	x_i	$x_i - 5$	$ x_i - 5 $	$(x_i - 5)^2$
x_1	3	-2	2	4
x_2	5	0	0	0
x_3	6	1	1	1
x_4	8	3	3	9
Σ		2	6	14

Angenommene Mittelwerte zwischen 1 und 10 (Ausschnittsvergrößerung um 5.5)



Nur die Quadrierung führt zu einem Minimum beim ‘wahren’ Mittelwert \Rightarrow ein-eindeutiger Zusammenhang zwischen Varianz und Mittelwert (aus den Daten und der Varianz kann man auf den Mittelwert schließen)

Summe der ‘Abweichungsquadrate’ als Abstandsmaß:

$$a^2 + b^2 = c^2 \quad \text{mit } a = (x_1 - \bar{x})^2 \quad \text{und} \quad b = (x_2 - \bar{x})^2$$

liefert c den kürzesten ‘Abstand’ durch die von a und b aufgespannte ‘Strecke’. Dies lässt sich auf höhere Dimensionen verallgemeinern ($a^2 + b^2 + c^2 + \dots = x^2$) – nichts anderes ist die

Summe der Abweichungsquadrate (SAQ) = Sum of Squares (SSq): $\sum (x_i - \bar{x})^2$

Mittlere Quadratsumme (MQS) = Mean Square (MS) = Varianz: $\sum (x_i - \bar{x})^2 / n$ (bzw. $\sum (x_i - \bar{x})^2 / (n-1)$)

Standardabweichung (St.Abw.) = Standard Deviation (sd) = $\sqrt{\text{Varianz}}$
(nicht ein-eindeutig, da positive und negative Wurzel existieren kann)

ANOVA - Terminologie:

Faktoren = factors

Ausprägungen = levels

Interaktionen = interactions = crossing (z.B. gender vs. vowels)

Nesting (z.B. frontness vs. vowels)

unklar (z.B. Plosive sth/stl in Dutch -- kein /g/)

Random factors (VPs / S_j, Probanden, Teilnehmer, Participants)

unbalanced designs

homogeneity

standard distribution

ANOVA

lm(F1~Gender

JMP (equal_f_m.jmp):

t-test:

Gender	df	Mean	Varianz
female	49	417.041	8,460.504361
male	49	521.612	37,778.141956

t-ratio = 3.40415 (unequal variance)

df= 68.47274

p = 0.0011

F-test:

Gender	df	SSq	MSq
Model	1	267912	267912
Error	96	2219453.6	23119
C.Total	97	2487365.6	

(overall data)

Mean = 469.32653

sd = 160.13414

variance = 25,642.9

SSq = 2,487,365.6

SSq-female: 406,104

SSq-male: 1,813,351

SSq of m+f = error: 2,219,455

F-ratio = 11.5882

df = (1, 96)

p = 0.001

gender_100

f, N= 363, mean = 516.972, sd=182.553

m, N=536, mean=528.188, sd=273.778

t-test Gender_100 male/female

t-ratio = 0.736928; DF=896.8043

p=0.4614

Anova (different due to missing value!!)

Model: DF: 1 SSq: 27226 MSq: 27226.2

Error: DF: 897 SSq: 52164344 MSq: 58154.2

Total: DF: 898 SSq: 52191570

Mean all data: 523.65962; sd: 241.08046

F-ratio = 0.4682

p=0.4940